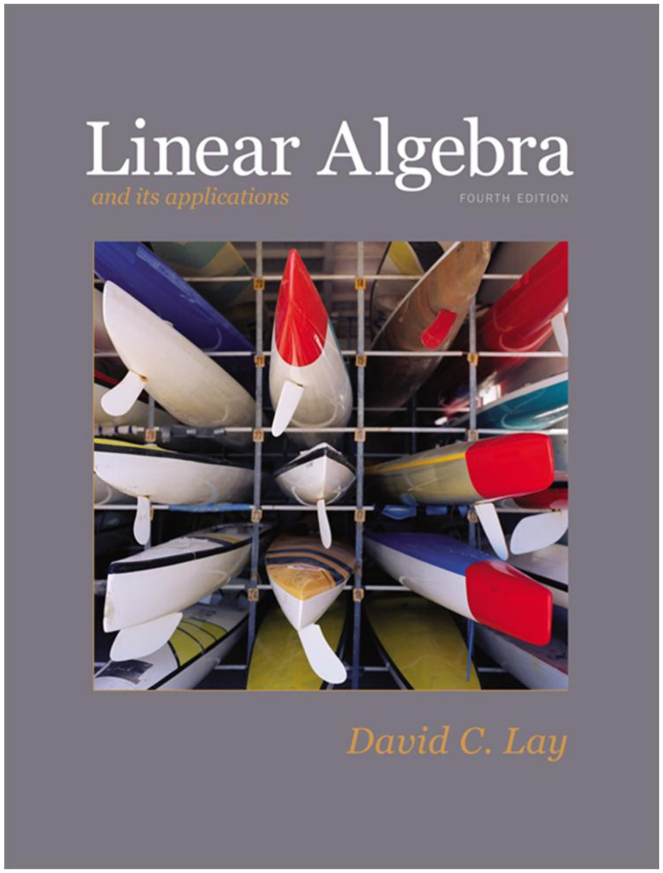
7 Symmetric Matrices and Quadratic Forms

7.2

QUADRATIC FORMS



QUADRATIC FORMS

- A **quadratic form** on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose value at a vector **x** in $\mathbb{R}^n$ can be computed by an expression of the form $Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x}$, where A is an $n \times n$ symmetric matrix.

- The matrix A is called the **matrix of the quadratic form**.

QUADRATIC FORMS

- **Example 1:** Let $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^T A\mathbf{x}$ for the following matrices.

  a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

  b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

QUADRATIC FORMS

- **Solution:**

a. $\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. There are two -2 entries in A.

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{bmatrix}$$

$$= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2)$$

$$= 3x_1^2 - 2x_1x_2 - 2x_2x_1 + 7x_2^2$$

$$= 3x_1^2 - 4x_1x_2 + 7x_2^2$$

QUADRATIC FORMS

- The presence of $-4x_1x_2$ in the quadratic form in Example 1(b) is due to the -2 entries off the diagonal in the matrix A.
- In contrast, the quadratic form associated with the diagonal matrix A in Example 1(a) has no x_1x_2 *cross-product* term.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- If **x** represents a variable vector in $\mathbb{R}^n$, then a **change of variable** is an equation of the form

$$\mathbf{x} = P\mathbf{y}, \text{ or equivalently, } \mathbf{y} = P^{-1}\mathbf{x} \qquad \text{----(1)}$$

  where P is an invertible matrix and **y** is a new variable vector in $\mathbb{R}^n$.

- Here **y** is the coordinate vector of **x** relative to the basis of $\mathbb{R}^n$ determined by the columns of P.
- If the change of variable (1) is made in a quadratic form $\mathbf{x}^T A\mathbf{x}$, then

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T (P^T AP)\mathbf{y} \quad \text{----(2)}$$

  and the new matrix of the quadratic form is P^TAP.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Since A is symmetric, Theorem 2 guarantees that there is an orthogonal matrix P such that P^TAP is a diagonal matrix D, and the quadratic form in (2) becomes $\mathbf{y}^TD\mathbf{y}$.

- **Example 2:** Make a change of variable that transforms the quadratic form $Q(\mathrm{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.

- **Solution:** The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

CHANGE OF VARIBALE IN A QUADRATIC FORM

- The first step is to orthogonally diagonalize A.
- Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$.
- Associated unit eigenvectors are

$$\lambda = 3: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \lambda = -7: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

- These vectors are automatically orthogonal (because they correspond to distinct eigenvalues) and so provide an orthonormal basis for $\mathbb{R}^2$.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Let

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

- Then $A = PDP^{-1}$ and $D = P^{-1}AP = P^{T}AP.$

- A suitable change of variable is

$$\mathrm{x} = P\mathrm{y}, \text{where } \mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Then

$$x_1^2 - 8x_1x_2 - 5x_2^2 = \mathrm{x}^T A\mathrm{x} = (P\mathrm{y})^T A(P\mathrm{y})$$
$$= \mathrm{y}^T P^T AP\mathrm{y} = \mathrm{y}^T D\mathrm{y}$$
$$= 3\mathrm{y}_1^2 - 7\mathrm{y}_2^2$$

- To illustrate the meaning of the equality of quadratic forms in Example 2, we can compute $Q(\mathbf{x})$ for $\mathrm{x} = (2, -2)$ using the new quadratic form.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- First, since $\mathrm{x} = P\mathrm{y}$,

$$\mathrm{y} = P^{-1}\mathrm{x} = P^{T}\mathrm{x}$$

so

$$y = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

- Hence

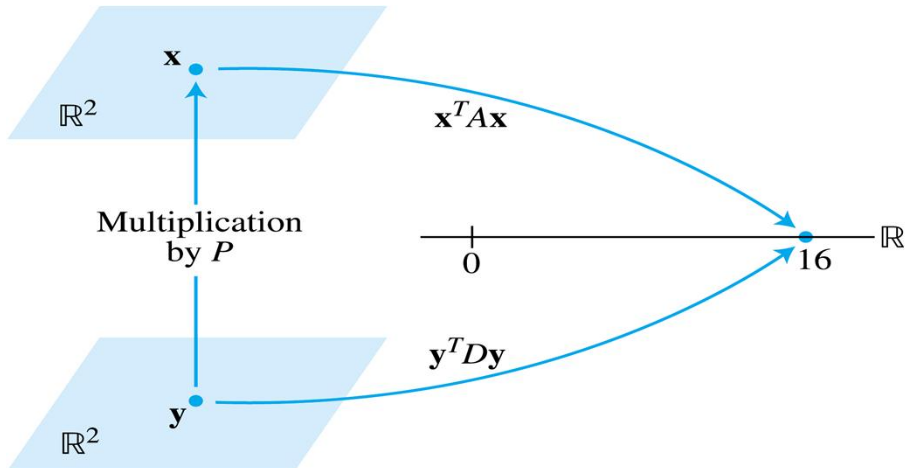
$$3\mathrm{y}_1^2 - 7\mathrm{y}_2^2 = 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5)$$

$$= 80/5 = 16$$

- This is the value of $Q(\mathbf{x})$ when $\mathrm{x} = (2, -2)$.

THE PRINCIPAL AXIS THEOREM

- See the figure below.



Change of variable in $\mathbf{x}^T A\mathbf{x}$.

- **Theorem 4:** Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathrm{x} = P\mathrm{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

THE PRINCIPAL AXIS THEOREM

- The columns of P in theorem 4 are called the **principal axes** of the quadratic form $\mathbf{x}^T A\mathbf{x}$.

- The vector **y** is the coordinate vector of **x** relative to the orthonormal basis of $\mathbb{R}^n$ given by these principal axes.

- **A Geometric View of Principal Axes**

- Suppose $Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x}$, where A is an invertible 2×2 symmetric matrix, and let c be a constant.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- It can be shown that the set of all **x** in $\mathbb{R}^2$ that satisfy

$$\mathbf{x}^T A \mathbf{x} = c \qquad \text{----(3)}$$

  either corresponds to an ellipse (or circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.

- If A is a diagonal matrix, the graph is in *standard position*, such as in the figure below.

$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1,\ a > b > 0$

ellipse

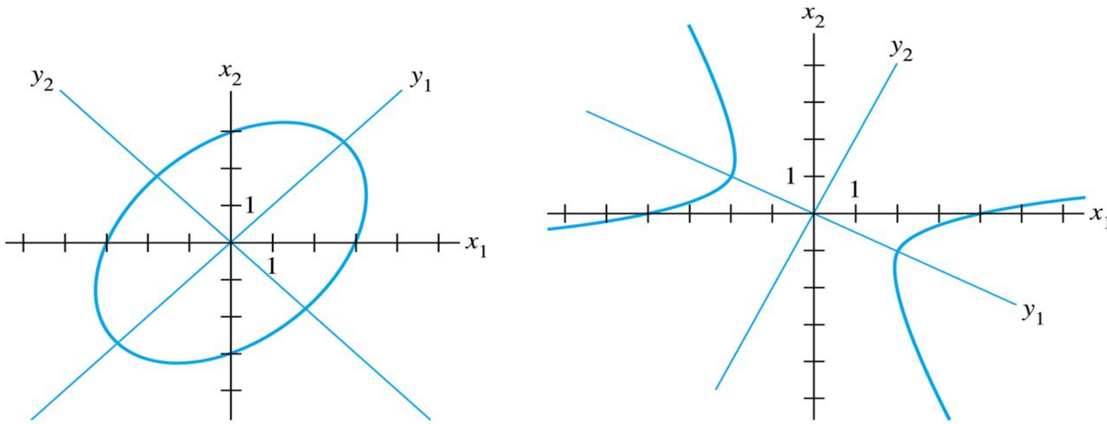
$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1,\ a > b > 0$

hyperbola

An ellipse and a hyperbola in standard position.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- If A is not a diagonal matrix, the graph of equation (3) is rotated out of standard position, as in the figure below.



(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$ (b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

An ellipse and a hyperbola *not* in standard position.

- Finding the *principal axes* (determined by the eigenvectors of A) amounts to finding a new coordinate system with respect to which the graph is in standard position.

CLASSIFYING QUADRATIC FORMS

- **Definition:** A quadratic form Q is:
  a. **positive definite** if $Q(\mathrm{x}) > 0$ for all $\mathrm{x} \neq 0$,
  b. **negative definite** if $Q(\mathrm{x}) < 0$ for all $\mathrm{x} \neq 0$,
  c. **indefinite** if Q (**x**) assumes both positive and negative values.

- Also, Q is said to be **positive semidefinite** if $Q(\mathrm{x}) \geq 0$ for all **x**, and **negative semidefinite** if $Q(\mathrm{x}) \leq 0$ for all **x**.

QUADRATIC FORMS AND EIGENVALUES

- **Theorem 5:** Let A be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}^T A\mathbf{x}$ is:
  a. positive definite if and only if the eigenvalues of A are all positive,
  b. negative definite if and only if the eigenvalues of A are all negative, or
  c. indefinite if and only if A has both positive and negative eigenvalues.

QUADRATIC FORMS AND EIGENVALUES

- **Proof:** By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathrm{x} = P\mathrm{y}$ such that

$$Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x} = \mathrm{y}^T D\mathrm{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \qquad \text{----(4)}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

- Since P is invertible, there is a one-to-one correspondence between all nonzero **x** and all nonzero **y**.

QUADRATIC FORMS AND EIGENVALUES

- Thus the values of $Q(\mathbf{x})$ for $\mathbf{x} \neq 0$ coincide with the values of the expression on the right side of (4), which is controlled by the signs of the eigenvalues $\lambda_1, \ldots, \lambda_n$, in three ways described in the theorem 5.